

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/116909>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# Intertemporal Price Discrimination with Two Products<sup>\*</sup>

Jean-Charles Rochet

SFI@UNIGE, University of Zürich and TSE

John Thanassoulis<sup>†</sup>

University of Warwick & CEPR<sup>‡</sup>

May 2, 2019

## Abstract

We study the two-product monopoly profit maximisation problem for a seller who can commit to a dynamic pricing strategy. We show that if consumers' valuations are not strongly-ordered then optimality for the seller can require intertemporal price discrimination: the seller offers a choice between supplying a complete bundle now, or delaying the supply of a component of that bundle until a later date. For general valuations we establish a sufficient condition for such dynamic pricing to be more profitable than mixed bundling. So we show that the Stokey (1979) no-discrimination-across-time result does not extend to two-product sellers when consumers' valuations are drawn from many standard distributions.

KEYWORDS: Multidimensional Mechanism Design; Second Degree Price Discrimination; Bundling; Time Discounting; Cross-sell; Substitutes and Complements.

JEL CLASSIFICATION: L11, D42

---

<sup>\*</sup>We would like to thank the editor, David Myatt, and three anonymous referees for invaluable guidance which has greatly improved this paper. We would also like to thank audiences at the CEPR IO 2015 workshop in Zürich, the Mid-west theory conference 2016 in Rochester, ESEM 2016 in Geneva, the Columbia-Duke-MIT-Northwestern IO Theory Conference 2016, the Paris meeting of CREST 2016, INFMA 2016 in Nanjing China, and the Berlin 2017 IO Day. We are very grateful for the helpful comments from Özlem Bedre-Defolie, Chris Dance, Sergiu Hart, Justin Johnson, Paul Klemperer, Jiangtao Li, Yossi Spiegel and Roland Strausz. Any errors remain our own.

<sup>†</sup>email: john.thanassoulis@wbs.ac.uk

<sup>‡</sup>Warwick Business School, University of Warwick.

# 1 Introduction

In celebrated work Stokey (1979) studied how a monopoly seller of a single good could change prices over time so as to maximise profits. It would seem natural that by lowering prices through time those who value the good most can be induced to buy the good earlier than those who value the good least, and so generate higher profits. Stokey (1979) demonstrated that for a seller who could commit to a path for prices and who had an inventory which could be restocked, such time varying prices were sub-optimal. The seller would do best by setting a price and sticking to it.

And yet marketing scholars have prominently argued that lowering prices over time to subsets of consumers, chosen based on purchase history, is profitable. Rossi et al. (1996) calculated that, twenty years ago, using purchase history to target price promotions could raise promotional revenues by a factor of two and a half times. More recently it has been noted that this is an under-estimate and the profit potential of personalising pricing based on prior purchase history appears to be greatest for online stores (Zhang and Wedel (2009), Taylor (2004)). However most marketing work on this issue has considered the question of who to target using purchase history data, and with what offers, separately (Prinzie and Van den Poel (2005), Reutterer et al. (2006)).

The purpose of this paper is to derive useful and interpretable sufficient conditions under which a seller's profit is increased by using prices which change over time, when the seller is not bound by capacity constraints and can commit to a path for prices.

We first study a two-good seller serving a consumer distribution with two types. We show that the optimal selling strategy hinges on whether or not the consumers' valuations are strongly-ordered. We define the valuations to *not* be strongly ordered if the two consumer types do not satisfy the standard Spence-Mirrlees sorting conditions: the consumer type valuing the bundle most does not also value each component most. We show that dynamic pricing is optimal in this case if and only if the consumers valuing the bundle most are numerous enough in the population. The most profit is achieved by using a strategy we describe as *dynamic pricing on the cross-sell*:<sup>1</sup> the seller offers a choice between supplying the complete bundle now, or delaying the supply of a component of that bundle until a later date. That is, the cross-sell to complete the bundle is delayed.

This rationalises a prominent sales approach in marketing. As an example, consider a company providing pay TV services. Let us suppose that type  $a$  consumers care little for sports on TV but value films & drama highly, while type  $b$  consumers value sports more than type  $a$  do, films & drama less than type  $a$ 's and yet value the overall TV package the most. This describes that consumer types  $a$  and  $b$  are not strongly ordered. Our work predicts that if the volume of sports-loving type  $b$  consumers is large enough then

---

<sup>1</sup>We borrow from marketing literature combining cross-selling and dynamic pricing. See Blattberg et al. (2008) and Ferrell and Hartline (2012) for textbook treatments.

the seller would optimally serve type  $b$  consumers with the bundle of sports and films & drama, while type  $a$  consumers would initially purchase just films & drama and receive targeted price offers which decline over time to add sports to their package; and they would purchase sports after delay.

Extending our analysis to continuous demand distributions is challenging. However here too we can offer a simple sufficient condition under which the strategy of dynamic pricing on the cross-sell is a profitable addition to mixed bundling prices. The sufficient condition is to establish whether or not the cross-partial derivative of the profit function with respect to the bundle price and a component price, evaluated at current mixed bundling prices, is negative. If it is then profits can be increased by offering a price reduction after some time for the cross-sell from the component to the whole bundle. And further we demonstrate that the sufficiency condition applies also when consumers have complementarities or substitutabilities in demand.

This partial derivative sufficiency condition is attractive in its simplicity and can be implemented analytically and numerically. Analytically we demonstrate that dynamic pricing on the cross-sell is a profitable addition to the best mixed bundling prices when consumers have valuations given by variations on the uniform distribution, exhausting most of the cases studied in the literature. Computationally we demonstrate that dynamic pricing on the cross-sell is more profitable if valuations are normally distributed with sufficient negative correlation.

The negative cross-partial condition allows for an economic intuition. The cross-partial of the profit function with respect to the bundle good and a component is a function of the density of consumers who are indifferent between buying the component and the bundle. A leading way in which this cross-partial becomes negative is if the density of consumers is downwards sloping across this group. An appropriate dynamic price for the cross-sell can be found which attracts just these consumers. The slope of the density function then ensures that more consumers upgrade to the cross-sell, than downgrade from the bundle to the delayed cross-sell, raising profits.

This paper has the following structure. We present the related literature in Section 2, and solve the discrete case as a motivating example in Section 3. The model (Section 4) and the analysis of the continuous case (Section 5) follow. The results for the continuous case are applied broadly in Section 6. Section 7 discusses real world examples, welfare, and technical robustness. Section 8 concludes with omitted proofs in the Appendix.

## 2 Literature Review

Our work builds on the celebrated result of Stokey (1979): that lowering prices over time to screen consumers would not be optimal for a single-product monopolist who could commit to prices. This is a surprising result, and since it was discovered scholars have

studied why we do see declining prices over time. An influential strand of research has dropped the assumption of seller commitment and has explored the implications if the seller is unable to commit to a path for prices. This led to the insight that such a seller would be driven to lower prices, potentially right down to the marginal cost of production (Coase (1972), Gul et al. (1986), Sobel (1991), Skreta (2006)). Our work studies optimal price dynamics when sellers have the ability to commit to a price path.

A parallel approach, influential in marketing and operations research, assumes that the seller has fixed capacity. This work is referred to as revenue management. Talluri and Van Ryzin (2006) provide a textbook treatment. Some of the most recent contributions in this area have assumed, as we do, that sellers have the ability to commit to a price path and buyers are rational in that they can alter the time at which they choose to buy (e.g. Besbes and Lobel (2015), Board and Skrzypacz (2016), Elmaghraby et al. (2008)). In these contributions the capacity constraints ensure that prices respond to the available inventory. Our work does not introduce capacity constraints.

Our work is distinguished from all of these strands of research by studying the sale of multiple products. The ability to link the price of a good to the purchase of other goods is known as bundling. Bundling is used in many retail markets such as cable TV operators bundling channels, banks bundling checking accounts with loan services and so on. Sub-additive bundling, the practice of offering a price reduction for buying multiple products, is optimal in general circumstances (Adams and Yellen (1976), McAfee et al. (1989), Armstrong (2013)). Our work builds on these results by deriving sufficient conditions under which profits can be improved even further by dynamic pricing.

The discount factor in the intertemporal pricing problem can be interpreted as a probability of delivery, an analogy which is not prominent in the literature, though we are not the first to make this link (Salant (1989)). Using this analogy we can translate what is known from the study of a monopolist selling lotteries back into the dynamic pricing problem studied here. The analogue of Stokey (1979) was demonstrated by Riley and Zeckhauser (1983): use of lotteries in sales are never optimal for a single good seller. Thanassoulis (2004) demonstrated that for a seller of substitutable goods, lotteries can be optimal. Our analysis confirms that delay can be part of the optimal selling mechanism for a seller of substitutable goods.

For the case of additive valuations without complementarities, Pavlov (2011a) solves fully the case in which consumers have valuations for two goods which are uniformly distributed on the square  $[\underline{x}, \underline{x} + 1]^2$ . His results, reinterpreted, imply that delay for some goods is optimal if  $\underline{x} > 0$ . Pavlov (2011b) considers more general distributions and his results imply that it cannot be optimal for a two good seller to delay the sale of both goods for a finite period of time. Hart and Reny (2015) study two specific examples and show that in these cases use of lotteries is optimal. Our work provides a general test for the profitability of delay, and analogously for the profitability of lotteries.

Solving for the optimal selling strategy requires one to solve a multidimensional screening problem, and it is known that this is often intractable (Rochet and Choné (1998)). By studying the dual problem Manelli and Vincent (2006) establish sufficient conditions under which stochastic pricing is suboptimal. However these conditions require the construction of a linear functional related to the dual problem and are therefore opaque. Daskalakis et al. (2017) link the duality approach to a field in operations research studying the transportation problem, and so derive a mechanism for solving for the optimal stochastic pricing schedule. The conditions for optimality are involved to apply.<sup>2</sup>

Determining implementable conditions which identify when intertemporal price discrimination is more profitable than fixed prices, which we seek to do here, is valuable for at least two reasons. Firstly, the interest from marketing scholars in the profits achievable has been discussed above. Secondly, on the theoretical side, it is possible for the profit gain in theory from using delay to be significant,<sup>3</sup> and yet identifying distributions for which the practice is even profitable remains unsolved.

Our work also sheds light on the study of sellers' optimal product lines. Note that a good delivered with delay is worth less to the consumer and so is a form of damaged good (Deneckere and McAfee (1996)). By using delay a seller is expanding her product line. Johnson and Myatt (2003) establish that if all products can be ranked in terms of quality, then for screening via multiple products to be optimal the surplus function must be log supermodular in quality and consumer type.<sup>4</sup> The standard way of defining consumer utility from delay of a given product naturally delivers a utility function which is not log-supermodular.<sup>5</sup> So, as per Stokey (1979), using delay is not optimal. But we show that with two products the profitability of creating multiple quality variants, e.g. by using delay, is re-established. This is not because of a property of the curvature of surplus, but because of the failure of a classic sorting condition which arises naturally when consumers differ in how they rank the products.

### 3 A motivating example

A seller produces two discrete goods denoted 1 and 2. There are two types of consumer,  $a$  and  $b$ . As we will consider delayed consumption of the bundle, it is important to capture the flow of utility provided by each good. Suppose that an  $a$  type of consumer derives flow utility from good 1 of  $ra_1$  per period of time where  $r$  is the discount rate. Immediate

---

<sup>2</sup>Our work contributes to the general endeavour, described above, to establish features of optimal selling mechanisms: for example Menicucci et al. (2015) on the optimality of pure bundling, and Tang and Wang (2017) on the optimal use of simple menus.

<sup>3</sup>Theorem A of Hart and Nisan (2017) re-interpreted as delay.

<sup>4</sup>With zero costs this implies that for  $q_2 > q_1$ ,  $u(x, q_2)/u(x, q_1)$  is strictly increasing in consumer type  $x$ . See also Anderson and Dana Jr (2009), Proposition 1.

<sup>5</sup>If the value to a consumer of type  $x$  of a good delayed by time  $t$  is  $xe^{-rt}$  where  $r$  is the discount rate, then setting quality to be  $q = e^{-rt}$ ,  $u(x, q) = xq$ . This utility is not strictly log-supermodular.

consumption of good 1 would therefore yield utility  $a_1 = \int_0^\infty r a_1 e^{-rt} dt$ . Similarly the utility from immediate consumption of good 2 is  $a_2$ , and that from the bundle is  $a_{12}$ . We allow the consumer to find the goods substitutes ( $a_{12} < a_1 + a_2$ ), or complements ( $a_{12} > a_1 + a_2$ ), or neither. Free disposal implies  $a_{12} \geq \{a_1, a_2\}$ . The notation for type  $b$  consumers is analogous. We denote the proportion of type  $a$  consumers  $\lambda^a$ ; the rest (proportion  $\lambda^b$ ) are type  $b$ . A consumer knows his or her type; the seller only knows the distribution of types in the population.

We label  $b$  the consumer type which values the bundle the most:

$$b_{12} > a_{12}. \quad (1)$$

Assuming bundle valuations grow in component valuations, we have that  $a$  consumers value at least one of the component goods less than type  $b$  consumers.<sup>6</sup> Without further loss of generality we label the indices so that

$$b_1 > a_1, \quad (2)$$

with strict inequality as we ignore the case of  $a$  and  $b$  being identical.

We assume the seller has marginal costs of production which equal zero, and the seller wishes to maximise her profits. The seller can produce according to demand and so capacity is not fixed. We assume that all consumers are present in the market at time  $t = 0$ , the seller can commit to a dynamic price path, and consumers have rational expectations.

Suppose a type  $a$  consumer receives good 2 before good 1, which she may or may not receive. Thus  $t_2^a \leq t_1^a$  capture the delivery times. The utility of the consumer is then

$$U(t_1^a, t_2^a; a) = \int_{t=t_2}^{t_1} r a_2 e^{-rt} dt + \int_{t=t_1}^{\infty} r a_{12} e^{-rt} dt - p^a,$$

where  $p^a$  is the  $t = 0$  net present value of all prices paid by  $a$ . Conducting the integration, setting  $q_i^a = e^{-rt_i^a}$  for  $i \in \{1, 2\}$  we have

$$U(q_1^a, q_2^a; a) = a_2 q_2^a + q_1^a (a_{12} - a_2) - p^a. \quad (3)$$

To demonstrate that intertemporal price discrimination is profitable in this setting, we consider the case in which consumers valuations are not strongly ordered. Thus even though  $b$  consumers value the bundle more than  $a$ ; type  $b$  consumers value good 2 less than  $a$ :

$$b_2 < a_2. \quad (4)$$

---

<sup>6</sup>In Section 4 we will formalise the value of the bundle and this assumption will remain respected.

It follows that the Spence-Mirrlees sorting conditions do not hold as  $a$  consumers value good 2 the most, (4), but then value the increment to the bundle the least:

$$(1) + (4) \Rightarrow a_{12} - a_2 < b_{12} - b_2. \quad (5)$$

Given the locations of the consumers identified by conditions (1), (2) and (4), if the seller were restricted to not price discriminate over time, it would be optimal to serve  $b$  consumers with the bundle, and type  $a$  consumers with the bundle or just good 2, depending on the valuations. In either case

$$t_1^b = t_2^b = t_2^a = 0 \quad \Rightarrow \quad q_1^b = q_2^b = q_2^a = 1. \quad (6)$$

Let us maintain that the goods in (6) are delivered immediately; even so intertemporal price discrimination can raise profits by setting  $q_1^a \notin \{0, 1\}$ .

The standard way to solve the seller's problem is to work just with the type  $a$  individual rationality constraint and the type  $b$  incentive compatibility constraint, optimise profits and then observe that the remaining constraints are satisfied. Proceeding in this manner and substituting (6) into (3) the seller's objective function is:

$$\begin{aligned} \max_{\{q_1^a, p^a, p^b\}} \Pi &:= \lambda^a p^a + \lambda^b p^b & (7) \\ \text{subject to} \quad a_2 + q_1^a(a_{12} - a_2) &\geq p^a & (IR_a) \\ b_{12} - p^b &\geq b_2 + q_1^a(b_{12} - b_2) - p^a & (IC_b) \end{aligned}$$

As profits increase in  $p^a$  it is optimal to raise the price paid by  $a$  consumers until  $(IR_a)$  is satisfied with equality. Similar reasoning yields that the price of good  $b$  should be raised so that  $(IC_b)$  is satisfied with equality:

$$p^a = a_2 + q_1^a(a_{12} - a_2) \quad , \quad p^b = (1 - q_1^a)(b_{12} - b_2) + p^a.$$

Substituting into the profit function we have

$$\Pi = a_2 + \lambda^b(b_{12} - b_2) + q_1^a [(a_{12} - a_2) - \lambda^b(b_{12} - b_2)]. \quad (8)$$

It is immediate that if  $\lambda^b$ , the proportion of type  $b$  consumers, is small enough then the coefficient of  $q_1^a$  is positive. In this case it would be optimal to set  $q_1^a = 1$  in which case both consumers receive the bundle immediately.

Given (5) if the proportion of type  $b$  consumers is large then the coefficient of  $q_1^a$  in (8) is negative. In this case optimality requires the delivery of good 1 to type  $a$  consumers to be delayed, so that  $q_1^a$  shrinks. Can it be delayed indefinitely so that type  $a$ 's do not receive good 1 and  $q_1^a = 0$ ? The answer is no, and the constraint arises from the individual



rationality constraint of type  $b$  consumers:

$$\begin{aligned}
IR_b : b_{12} &\geq p^b \\
&\Rightarrow b_{12} \geq (b_{12} - b_2)(1 - q_1^a) + a_2 + q_1^a(a_{12} - a_2) \\
&\Rightarrow q_1^a \geq \frac{a_2 - b_2}{a_2 - b_2 + b_{12} - a_{12}} .
\end{aligned} \tag{9}$$

Hence if  $\lambda^b$  is large enough, then the delivery time for good 1 to  $a$  is delayed to the point that  $q_1^a$  equals the lower bound in (9).<sup>7</sup> If the seller delayed the delivery of good 1 to  $a$  further, whilst keeping  $b$  just incentive compatible and holding  $a$  to zero utility, then  $b$  would receive negative utility and would not participate.

Finally we can confirm that the remaining constraint, the incentive compatibility constraint of type  $a$ , is indeed satisfied.<sup>8</sup>

And so the working above proves:<sup>9</sup>

**Proposition 1** *Suppose the two consumer types  $a$  and  $b$  are not strongly ordered so that (1), (2) and (4) hold. Suppose the seller delivers the bundle immediately to consumers  $b$  and good 2 immediately to consumers  $a$ . If there are enough type  $b$  consumers:*

$$\lambda^b > \frac{a_{12} - a_2}{b_{12} - b_2} \tag{10}$$

*then price discriminating over the delivery time of good 1 to type  $a$  consumers is more profitable than not. And in this setting good 1 is delivered most profitably to type  $a$  consumers at time*

$$t_1^a = \frac{1}{r} \ln \left( 1 + \frac{b_{12} - a_{12}}{a_2 - b_2} \right). \tag{11}$$

It is perhaps interesting to note that the proof above can be extended to show that the price discrimination of Proposition 1 is in fact fully optimal. It is not necessary to demonstrate this to show that the Stokey (1979) result does not extend to the multiple good setting, and so we relegate the details to Appendix A.1.

A profitable sales strategy in this two-type case can be achieved by offering a dynamic price only on the cross-sell. That is type  $a$  consumers can be sold good 2 immediately, and after time  $t_1^a$  given in (11) be offered the cross-sell of good 1 to complete their bundle at an overall price which is lower (in net present value terms) than the initial bundle price

---

<sup>7</sup>The lower bound in (9) is positive given (4) and (5).

<sup>8</sup>The  $IC_a$  constraint requires

$$0 \geq a_{12} - p^b = (1 - q_1^a)[(a_{12} - a_2) - (b_{12} - b_2)],$$

and this follows from (5).

<sup>9</sup>The development of Proposition 1 has been greatly improved by suggestions from the editor, David Myatt. Since solving Proposition 1 we have discovered that the special case of no complementarities in demand ( $a_{12} = a_1 + a_2$  and similarly for  $b$ ) is a corollary of some insightful earlier analysis contained in the working paper Pycia (2006).

which types  $b$  paid. We will explore this selling strategy in the more challenging setting of continuous consumer distributions in Section 5. Before then we offer some observations on the result of Proposition 1.

Note that Proposition 1 identifies as key a distribution of consumer types which is not strongly ordered, so that the standard sorting conditions do not hold: thus consumers of type  $a$  value good 2 more than types  $b$  do, whilst consumers of type  $b$  value good 1 more, and value the bundle more than the type  $a$ 's do. If consumers are strongly ordered then the insights pioneered by Stokey (1979) apply: finite delay is not optimal, consumers either buy a good immediately or are not served at all. The setting of not strongly-ordered consumers of multiple products arises naturally in many economically relevant contexts. The example of a seller providing pay-TV services, in which type  $a$  consumers liked films & drama while type  $b$ 's preferred sport, was presented in the Introduction.<sup>10</sup>

Reflection on Proposition 1 yields that price discrimination over time can be profitable in the one-good setting also, where there are multiple versions or qualities available. Relabel good 2 as the base good and good 1 as the luxury add-on then Proposition 1 yields conditions such that delaying the delivery of the add-on to  $a$  consumers, or otherwise degrading this add-on for  $a$  consumers, is more profitable than offering only immediate delivery. However in this add-ons interpretation type  $b$  consumers value the “no frills” base good less than  $a$ , but the version with luxury more than  $a$ . This may seem less natural.<sup>11</sup> This distribution of tastes would seem to arise more naturally, however, under the two good formulation we study here.

Next observe that the profitability of dynamic pricing on the cross-sell to complete the bundle to  $a$  requires type  $b$  consumers to be sufficiently numerous in the population. In this case it is particularly costly to the seller to leave a rent to  $b$  consumers. The tension arises as the type  $b$  consumers are the ones who value the bundle of both goods the most; hence to extract all of their rent with static prices would require bundle prices to be set so high that type  $a$  consumers would be, at least partially, unserved. Dynamic pricing allows a compromise to be found in which the seller simultaneously extracts the surplus of the type  $b$  consumers as well as that of the type  $a$  consumers. Using dynamic pricing for the component valued less by type  $a$  is the optimal way in which this can be accomplished.

Finally we consider how the extent of complementarities in demand alter the most profitable delivery time for the cross-sell (11). Holding component good valuations constant, the cross-sell becomes later the larger the gap between the bundle valuations of the two consumers:  $b_{12} - a_{12}$ . Thus if increasing the complementarity between the component goods causes the high type consumer's bundle valuation to grow more than the low type

---

<sup>10</sup>The requirement for valuations to be not strongly-ordered clearly has parallels with negatively correlated valuations. In such settings linking prices across products was shown in now classical work to increase the seller's profits (Adams and Yellen (1976), Schmalensee (1984)).

<sup>11</sup>Though it would follow if the high type  $b$  consumers had a better (unmodelled) outside option if denied the luxury version.

consumer's bundle valuation, then the cross-sell moves out in time. Price discrimination requires the utility from the lower quality option to be low enough to prevent high type consumers (here  $b$ ) deviating to choose it. As complementarities rise the delayed bundle at a price which  $a$  is willing to pay becomes increasingly attractive to  $b$ . To maintain the discrimination the delay option must be damaged further, and this is achieved by pushing out the point at which the reduction on the cross-sell price becomes available.

## 4 The Model

We develop the two type consumer case of Section 3.

A seller produces two discrete goods denoted by  $i$ , with  $i \in \{1, 2\}$ . There exists a population of consumers, each characterised by their type:  $x \triangleq (x_1, x_2) \in \mathbb{R}^2$ . As above a consumer with type  $x$  is assumed to derive flow utility from good  $i$  of  $rx_i$  per period of time where  $r$  is the discount rate, so  $x_i$  is the consumer's valuation for immediate consumption of a single unit of good  $i$ . We set the utility of the bundle to be

$$x_{12} := \gamma(x_1 + x_2).$$

The parameter  $\gamma$  is assumed to be the same for all consumers and we restrict to  $\gamma \in (1/2, \infty)$ . If  $\gamma > 1$  then we are modelling complements, whereas  $\gamma \in (1/2, 1)$  captures substitutes. This approach to substitutes and complements is tractable and has been used in the literature: Venkatesh and Kamakura (2003), Bakos and Brynjolfsson (1999). The literature on randomness and pricing (e.g. Manelli and Vincent (2006)) has typically restricted analysis to the case in which the goods are neither substitutes nor complements, so that utility for the bundle is strictly-additive in the component valuations; that is  $\gamma = 1$ .

The seller has zero marginal cost and wishes to maximise her profits. To achieve this the seller is able to bundle the products and so can offer the menu of take-it-or-leave-it prices  $\{p_1, p_2, p_B\}$  with  $p_B$  being the price for the bundle of both goods. We refer to these prices as the bundling prices.<sup>12</sup> In addition to the bundling prices we allow the seller to offer dynamic prices. In part inspired by Proposition 1, we will derive testable conditions guaranteeing that profit increases can be achieved by using a menu of cross-sell prices such that if a consumer buys component good 2, then after a time delay of  $\tau$  the same consumer may also buy good 1 to complete her bundle for a cross-sell price of  $p_{1X}(\tau)$ . As noted above, we refer to this as dynamic pricing on the cross-sell.

We reiterate that production of each unit occurs at the time of the consumer's consumption, so the seller does not have a fixed capacity to sell; that all consumers are

---

<sup>12</sup>In some works these prices are referred to as mixed-bundling prices to distinguish them from pure-bundling in which only the bundle of both goods is sold. We drop the prefix 'mixed' for expositional clarity where appropriate.

present in the market at time  $t = 0$ ; that the seller can commit to a dynamic price path; and that consumers have rational expectations.

Suppose consumer  $x$  receives good  $i$  at time  $t_i$  and pays a net price  $p$ . Setting  $(q_1, q_2) = (e^{-rt_1}, e^{-rt_2})$ , we rework (3) as described in Appendix A.1 to establish the consumer's utility:

$$U(q; x) = q_1 x_1 + q_2 x_2 + (\gamma - 1) \min(q_1, q_2) \cdot (x_1 + x_2) - p. \quad (12)$$

Each consumer knows their valuation vector, but the seller only knows the distribution of valuations in the population,  $f(x_1, x_2)$ . We assume the density function is differentiable with a finite mean; the latter assumption implies that the resulting profit function has a maximum.

## 5 Increasing Profits with Dynamic Pricing – The Continuous Case

The fully optimal multidimensional screening contract is difficult to establish (Rochet and Choné (1998)). We pursue a more modest goal: to derive a (useful) sufficiency condition as to when dynamic pricing can raise profits beyond bundling prices. Let us suppose that the seller of our model has set prices  $\{p_1, p_2, p_B\}$ . We do not require these prices to be optimal amongst the set of all bundling prices. The work of this section will apply more generally. We will study optimal bundling prices subsequently.

We assume consumers can choose to purchase anonymously and therefore it follows that the bundling prices must be sub-additive:

$$p_B \leq p_1 + p_2. \quad (13)$$

We make the following assumption:

$$\textbf{Assumption: } p_B < \gamma(p_1 + p_2). \quad (14)$$

Condition (14) is immediate from (13) if the goods are complements ( $\gamma > 1$ ). If  $\gamma = 1$ , so goods are neither complements nor substitutes, condition (14) only requires the seller to have introduced a bundle price reduction; if valuations are separable then this is required for optimality (McAfee et al. (1989)). Only if the goods are substitutes ( $\gamma < 1$ ) does the assumption have bite. Consumers will prefer the bundle to not purchasing if  $\gamma(x_1 + x_2) - p_B \geq 0$ . Thus condition (14) ensures there exist valuations at which consumers are indifferent between buying the bundle and not participating. The assignment of consumers to products under bundling prices is depicted graphically in Figure 1.

The total profit of the seller  $\pi$  is the sum of profits from the sale of good 1, good 2

and the bundle which we denote

$$\pi = \pi_1(p_1, p_B) + \pi_2(p_2, p_B) + \pi_B(p_1, p_2, p_B). \quad (15)$$

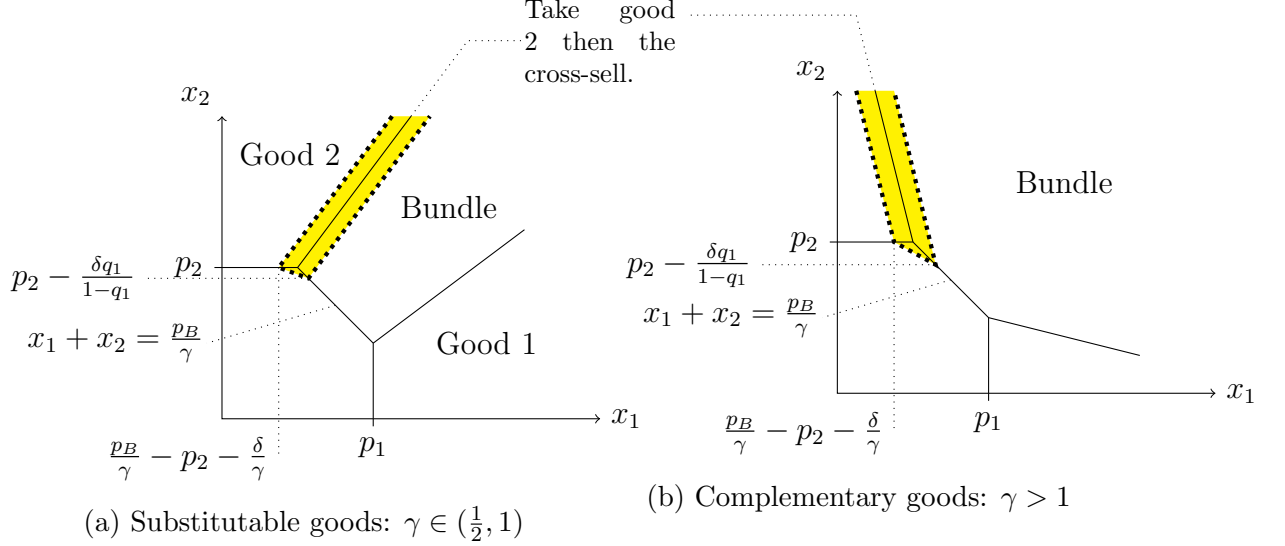


Figure 1: The participation regions under complementarities in demand

Notes: When the cross-sell of good 1 is offered to good 2 purchasers at time  $T$  for a price of  $p_B - p_2 - \delta$ , and setting  $q_1 = e^{-rT}$ , then the shaded band accept.

Using Figure 1:

$$\begin{aligned} \pi_1(p_1, p_B) &= \int_{x_1=p_1}^{\infty} \int_{x_2=0}^{\frac{1-\gamma}{\gamma}x_1 + \frac{1}{\gamma}(p_B-p_1)} p_1 dF, & \pi_2(p_2, p_B) &= \int_{x_2=p_2}^{\infty} \int_{x_1=0}^{\frac{1-\gamma}{\gamma}x_2 + \frac{1}{\gamma}(p_B-p_2)} p_2 dF, \\ \pi_B(p_1, p_2, p_B) &= p_B \left\{ \int_{x_2=p_2}^{\infty} \int_{x_1=\frac{1-\gamma}{\gamma}x_2 + \frac{1}{\gamma}(p_B-p_2)}^{x_2+p_1-p_2} dF + \int_{x_1=p_1}^{\infty} \int_{x_2=\frac{1-\gamma}{\gamma}x_1 + \frac{1}{\gamma}(p_B-p_1)}^{x_1-p_1+p_2} dF \right. \\ &\quad \left. + \int_{x_2=\frac{p_B}{\gamma}-p_1}^{p_2} \int_{x_1=\frac{p_B}{\gamma}-x_2}^{p_1} dF \right\}. \end{aligned} \quad (16)$$

We restrict attention to the case in which all product combinations are purchased by a positive measure of consumers, so that none of the prices posted are unused by consumers. Following Manelli and Vincent (2006) we refer to this assumption as **ABS** – All Bundles Sold.

**ABS** At the posted bundling prices, non-zero volumes of all product combinations are sold.

Assumption **ABS** guarantees that the consumer with valuation  $\left(\frac{p_B}{\gamma} - p_2, p_2\right)$  lies strictly within the support of the consumer type space. Under general distributions of consumer valuations it is a weak assumption. However many analytically studied cases, such as variations of the uniform distribution on the unit-square, heavily restrict the

support of the density function. In such cases assumption ABS has more bite. We will explore the ABS assumption in more detail in Section 7.3.

We can now present this section's main result:

**Proposition 2** *Suppose that a seller can lower prices at a fixed time  $T > 0$  after the initial prices  $\{p_1, p_2, p_B\}$  are posted. Suppose bundling prices satisfy condition (14) and ABS. If in the absence of dynamic pricing*

$$\left. \frac{\partial^2 \pi}{\partial p_2 \partial p_B} \right|_{(p_1, p_2, p_B)} < 0, \quad (\text{X-sell})$$

*then seller profits can be increased further by offering a dynamic cross-sell such that buyers of good 2 may purchase the cross-sell of good 1 at time  $T$  at a reduced cross-sell price of  $p_{1X} = p_B - p_2 - \delta$ , for small  $\delta$ .*

**Proof.** Setting  $q_1 = e^{-rT}$  and  $q_2 = 1$ , we invoke the utility function (12) to establish that a consumer  $x$  will prefer to buy good 2 followed by the cross-sell to buying the bundle of both goods immediately if

$$\gamma(x_1 + x_2) - p_B < q_1 x_1 + x_2 + (\gamma - 1)q_1(x_1 + x_2) - (p_2 + q_1 p_{1X}).$$

Substituting for the cross-sell price and simplifying yields:

$$x_1 < \frac{1 - \gamma}{\gamma} x_2 + \frac{1}{\gamma} (p_B - p_2) + \frac{q_1}{1 - q_1} \frac{\delta}{\gamma}.$$

Proceeding similarly we see that the introduction of the cross-sell alters the consumer participation regions in the shaded areas of Figure 1.

The sale of good 2 followed by the cross sell at time  $T$  generates a profit of

$$p_2 + e^{-rT} p_{1X} = q_1 p_B + (1 - q_1) p_2 - q_1 \delta.$$

Using Figure 1 we can determine that the introduction of the cross-sell offer changes seller profits by

$$\begin{aligned}
\Delta\Pi = & [q_1 p_B + (1 - q_1) p_2 - q_1 \delta] \left\{ \int_{x_2=p_2-\delta\frac{q_1}{1-q_1}}^{p_2} \int_{x_1=-x_2\left(\frac{1-q_1+q_1\gamma}{q_1\gamma}\right)+\frac{p_B}{\gamma}+\left(\frac{1-q_1}{q_1\gamma}\right)p_2-\frac{\delta}{\gamma}}^{\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}+\frac{\delta}{\gamma}\frac{q_1}{1-q_1}} dF \right. \\
& \left. + \int_{x_2=p_2}^{\infty} \int_{x_1=\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}-\frac{\delta}{\gamma}}^{\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}+\frac{\delta}{\gamma}\frac{q_1}{1-q_1}} dF \right\} \\
& - p_2 \int_{x_2=p_2}^{\infty} \int_{x_1=\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}-\frac{\delta}{\gamma}}^{\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}} dF \\
& - p_B \left\{ \int_{x_2=p_2}^{\infty} \int_{x_1=\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}}^{\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}+\frac{\delta}{\gamma}\frac{q_1}{1-q_1}} dF \right. \\
& \left. + \int_{x_2=p_2-\delta\frac{q_1}{1-q_1}}^{p_2} \int_{x_1=-x_2+\frac{p_B}{\gamma}}^{\left(\frac{1-\gamma}{\gamma}\right)x_2+\frac{p_B-p_2}{\gamma}+\frac{\delta}{\gamma}\frac{q_1}{1-q_1}} dF \right\}. \tag{17}
\end{aligned}$$

It is apparent from Figure 1 that small price reductions  $\delta$ , on the cross-sell at time  $T$  lead to first order gains from some good 2 consumers buying the cross-sell, and first order losses from the price reduction to bundle consumers. One can show that these effects cancel:  $\lim_{\delta \searrow 0} \frac{d}{d\delta} \Delta\Pi = 0$ . This implies that a more nuanced analysis is required.

We prove in Lemma 2, contained in the Appendix, that  $\lim_{\delta \searrow 0} \frac{d^2}{d\delta^2} \Delta\Pi > 0$  if and only if (X-sell) holds. This then proves the result. ■

In Section 6 we will show that using Proposition 2 we can demonstrate that dynamic pricing of the cross-sell increases profits beyond the best mixed bundling prices in a wide range of situations. Shortly we will explore the intuition as to why the cross-derivative of the profit function, given as (X-sell), is key to the profitability of dynamic pricing. First we note that Proposition 2 considers just one sort of dynamic pricing: consumers may purchase one good at time 0, and the price to complete the bundle is lowered after a period of time. In the case of no complementarities in demand ( $\gamma = 1$ ) we know from Pavlov (2011b) that it cannot be optimal to have positive delay for both components. The inspiration behind focusing on the case in which delay is on the cross-sell price reduction comes from our analysis of the discrete case in Section 3. There we demonstrated that dynamic pricing of the cross-sell is the unique form of dynamic pricing which can form part of the optimal selling strategy in that setting.

The neatness of (X-sell) suggests that a deeper intuition should exist. We try to outline this intuition now. Suppose the seller has set bundling prices  $\{p_1, p_2, p_B\}$ . The purchase choices can be depicted as in Figure 2. Let us denote the mass of bundle consumers by  $\mathcal{B}$  and use  $\mathcal{I}_{2B}$  to denote the line integral along the boundary between the purchases of good 2 and the bundle;  $\mathcal{I}_{1B}$  to denote the line integral along the boundary between the purchases of good 1 and the bundle; and  $\mathcal{I}_{\emptyset B}$  to denote the line integral along the boundary between those who do not purchase and the bundle purchasers. The variables are depicted in Figure 2. We can therefore express the first derivative of profit as:

$$\frac{\partial \pi}{\partial p_B} = \mathcal{B} - (p_B - p_2) \mathcal{I}_{2B} - p_B \mathcal{I}_{\emptyset B} - (p_B - p_1) \mathcal{I}_{1B}. \tag{18}$$

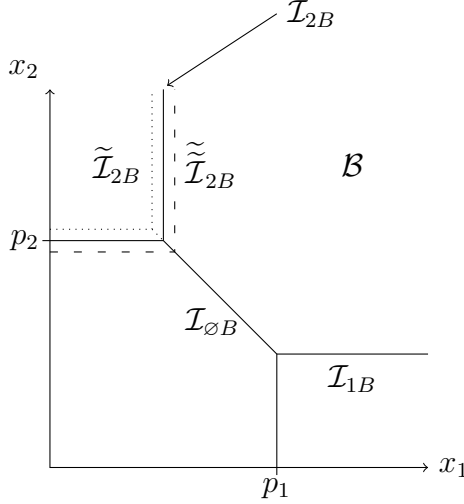


Figure 2: Graphical depiction of the cross-partial of the profit function

Now consider raising the price of good 2 by a small amount holding the bundle price constant. This alters the boundaries of the good 2 purchases to the dotted lines in Figure 2. Suppose condition (X-sell) of Proposition 2 holds:  $\frac{\partial}{\partial p_2} \frac{\partial \pi}{\partial p_B} < 0$ . Thus raising  $p_2$  slightly must lower  $\frac{\partial \pi}{\partial p_B}$  yielding

$$\frac{\partial \pi}{\partial p_B}(p_1, p_2 + \varepsilon, p_B) < \frac{\partial \pi}{\partial p_B}(p_1, p_2, p_B) \text{ for small enough } \varepsilon > 0. \quad (19)$$

Using Figure 2, and analogously to (18), the first derivative of the profit function with respect to  $p_B$  at prices  $(p_1, p_2 + \varepsilon, p_B)$  is, to first order in  $\varepsilon$ ,

$$\frac{\partial \pi}{\partial p_B}(p_1, p_2 + \varepsilon, p_B) \approx \mathcal{B} + \varepsilon \mathcal{I}_{2B} - (p_B - p_2 - \varepsilon) \tilde{\tilde{\mathcal{I}}}_{2B} - p_B \mathcal{I}_{\emptyset B} - (p_B - p_1) \mathcal{I}_{1B}. \quad (20)$$

Where we have used the fact that at prices  $(p_1, p_2 + \varepsilon, p_B)$  the mass of bundle purchasers is, to first order in  $\varepsilon$ ,  $\mathcal{B} + \varepsilon \mathcal{I}_{2B}$ . Combining (18) and (20) in the inequality (19) yields  $\tilde{\tilde{\mathcal{I}}}_{2B} > \left( \frac{p_B - p_2 + \varepsilon}{p_B - p_2 - \varepsilon} \right) \mathcal{I}_{2B} > \mathcal{I}_{2B}$ . A similar thought experiment for the case of lowering  $p_2$  would deliver that  $\tilde{\tilde{\mathcal{I}}}_{2B} > \mathcal{I}_{2B} > \tilde{\mathcal{I}}_{2B}$ . Thus we see that the cross-partial condition (X-sell) is related to the slope of the density function across the participation boundary between the purchasers of good 2 and the bundle.

Now consider the introduction of the dynamic cross-sell at time  $T$  studied in Proposition 2. We know from Figure 1 that this cross-sell offer allows the seller to target the boundary  $\mathcal{I}_{2B}$  and not the other boundaries of the purchase regions. Consumers between  $\tilde{\tilde{\mathcal{I}}}_{2B}$  and  $\mathcal{I}_{2B}$  add good 1 bought with delay to their purchase of good 2 and so pay more. This raises profits by approximately  $\tilde{\tilde{\mathcal{I}}}_{2B} \cdot e^{-rT} (p_B - p_2)$ . However consumers between  $\mathcal{I}_{2B}$  and  $\tilde{\mathcal{I}}_{2B}$  now delay the creation of profit  $(p_B - p_2)$  for time  $T$  and so profits drop by approximately  $\tilde{\mathcal{I}}_{2B} \cdot e^{-rT} (p_B - p_2)$ . As condition (X-sell) implies that  $\tilde{\mathcal{I}}_{2B} > \tilde{\tilde{\mathcal{I}}}_{2B}$  overall



this is a positive contribution to profits.

Summing up therefore, the negative cross-partial in the profit function (X-sell) implies a downwards sloping density function across the boundary between purchasers of good 2 and the bundle. Thus there is a build up of consumers of good 2 who nearly buy the bundle; and there is a comparatively small mass of bundle consumers at risk of reducing their purchases to just good 2. Combining these two groups through dynamic pricing on the cross-sell is profitable as more consumers trade-up to, rather than trade-down from, the bundle. We will see this insight at work in the next section.

## 6 Continuous Distributions – Examples

Proposition 2 is readily applied to mixed bundling prices, and so we have a testable sufficient condition for the profitability of intertemporal price discrimination over and above mixed bundling. In this section we consider a variety of continuous distributions which are prominent in the theoretical and empirical literature and so show straightforwardly that dynamic pricing on the cross-sell is profitable generally. In the examples that follow we will restrict attention to the setting in which the utility of the bundle is strictly additive,  $\gamma = 1$ , implying the goods are neither complements nor substitutes.

### 6.1 The Multivariate Normal Distribution

First we study consumer valuation density functions given by a multivariate normal distribution. Schmalensee (1984) was one of the first authors to explore this setting which he identified as important given the “*frequency with which normal distributions arise in the social sciences*”.<sup>13</sup> Schmalensee (1984) used empirical techniques to establish conditions under which mixed bundling generates strictly more profits than pure bundling or pure component prices. In this section we study the distribution of consumer tastes:

$$x \sim N \left( \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right). \quad (21)$$

Thus consumer valuations have mean  $(2, 2)$  and correlation-coefficient  $\rho$ . For a given  $\rho$  we follow the techniques pioneered by Schmalensee (1984) and solve for the best mixed bundling prices. This is done by numerically evaluating the first order conditions given by differentiating (16).<sup>14</sup> The best mixed bundling prices for a subset of  $\rho$  values are given in Table 1. Our sufficiency condition can now be implemented.

---

<sup>13</sup>Schmalensee (1984) page S212.

<sup>14</sup>This step uses optimisation routines contained in standard software. The library optimisation routine in Matlab we use implements the simplex search method of Lagarias et al. (1998).

Correlation coefficient ( $\rho$ )	component price( $p$ )	bundle price ( $p_B$ )
0.5	2.134	3.264
0.25	2.285	3.226
0	2.432	3.187
-0.25	2.588	3.153
-0.5	2.773	3.139
-0.75	3.043	3.189

Table 1: Optimal mixed bundling prices with normally distributed consumer tastes according to (21).

**Claim 1** *Suppose consumer valuations are distributed according to the normal distribution given in (21), then profits can be increased beyond that available from optimal mixed bundling by offering a delayed price reduction on the cross-sell if  $\rho \leq -0.47$ .*

**Proof.** The proof is numerical. First, for given  $\rho$  evaluate the best bundling prices (some of which are depicted in Table 1), and then evaluate condition (X-sell).<sup>15</sup> Repeating for multiple values of  $\rho$  delivers Figure 3, which depicts the cross-partial derivative,  $\partial^2\pi/\partial p_2\partial p_B$ , evaluated at the best bundling prices for given  $\rho$ . The result then follows by inspection. ■

Negative correlation has, as noted above, long been associated with the optimality of bundling over pure component prices. Claim 1 demonstrates that negative correlation is also important for the profitability of dynamic pricing on the cross-sell.

## 6.2 Independently Distributed Valuations

Much theoretical work has focused on special cases in which valuations are independently distributed across the goods; the leading one being valuations distributed uniformly on the unit-square. An early contribution in this setting is due to Manelli and Vincent (2006), Theorem 4, which proves that mixed bundling with no delayed discounting is fully optimal if:<sup>16</sup>

$$\frac{xf'_i(x)}{f_i(x)} \text{ is increasing in } x \ \forall x \in \text{supp}f_i \text{ and all } i \in \{1, 2\}. \quad (\text{MV})$$

Thus (MV) informs us that a sufficient condition to guarantee that no delay is optimal is that the elasticity of the density function is increasing for both goods. We can identify an alternative sufficiency condition under which delayed discounting is more profitable than any mixed bundling prices.

<sup>15</sup>Formally we use the expression for the cross-partial derivative given in the proof of Proposition 2. A local linearisation is used to approximate for the first derivative term arbitrarily closely.

<sup>16</sup>Manelli and Vincent (2006) also maintain three assumptions: that the density is supported on  $[0, 1]^n$ ; is independent across variables; and satisfies the condition  $f(x_1, \dots, x_n) + x_i\partial f/\partial x_i \geq 0, \forall x\forall i$ .

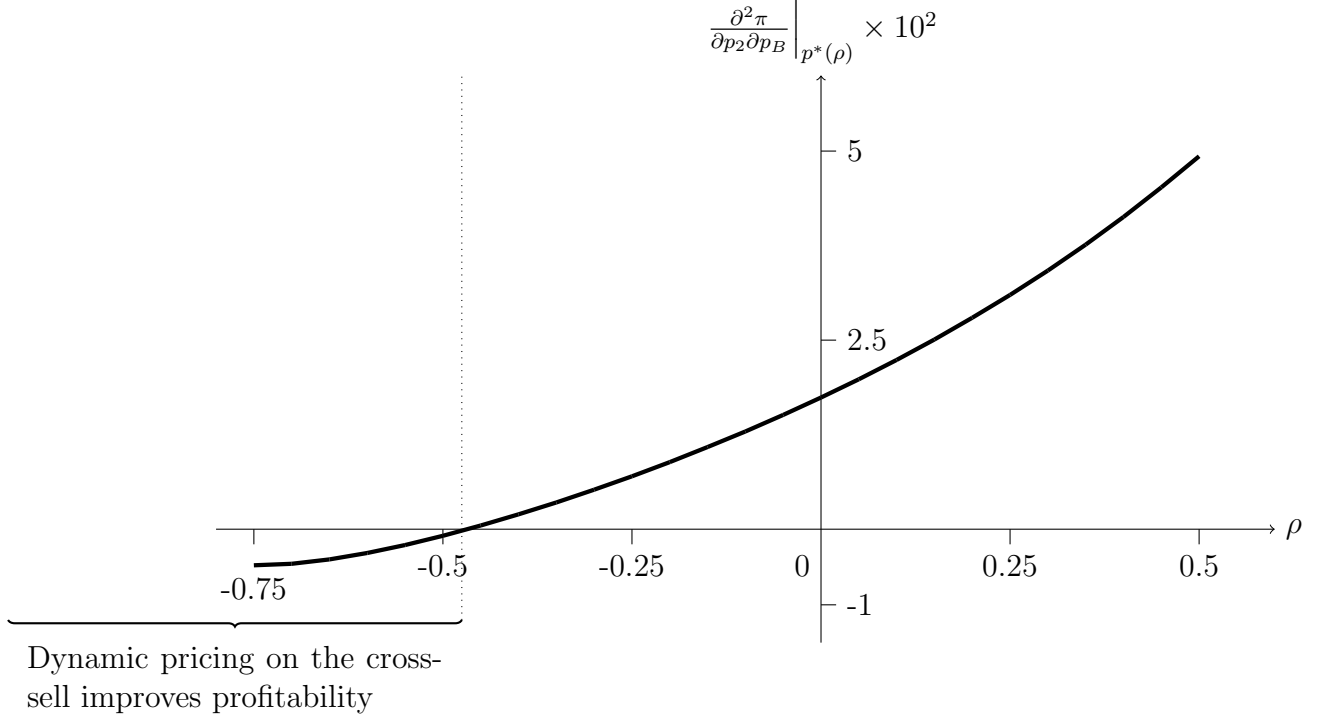


Figure 3: Cross-partial of profit with normally distributed valuations

Notes: The optimal bundling prices ( $p^*(\rho)$ ) are calculated assuming that consumer valuations are distributed according to the bivariate normal distribution with mean  $\mu = (2, 2)$ , variance-covariance matrix  $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ , and marginal costs of production are normalised to zero.  $\rho$  is therefore the correlation coefficient between the valuations of goods 1 and 2. The cross-partial derivative,  $\partial^2 \pi / \partial p_2 \partial p_B$ , is evaluated at the optimal bundling prices and its value is plotted above. Proposition 2 shows that profits are increased by the introduction of dynamic pricing on the cross-sell when  $\partial^2 \pi / \partial p_2 \partial p_B < 0$ .

**Corollary 1** *Suppose consumer valuations are independent across goods so that  $f(x) = f_1(x_1) \times f_2(x_2)$ , and that consumers see the goods as neither complements nor substitutes ( $\gamma = 1$ ). If at the best bundling prices assumption ABS holds, then dynamic pricing on the cross-sell of good 1 at time  $T$  increases profits if*

$$\frac{x_1 f_1(x_1)}{F_1(x_1)} \text{ is declining in } x_1 \text{ at } x_1 = p_B - p_2 \in \text{supp} f_1. \quad (\text{RT})$$

Observe that if (RT) holds for all  $x_1 \in \text{supp} f_1$  then it holds at  $x_1 = p_B - p_2$ . This latter formulation is expressed entirely in terms of model fundamentals. Thus (RT) yields that delayed discounting is profitable if the elasticity of the distribution is declining for at least one of the goods. In Section 6.2.3 we will explore how the two conditions (MV) and (RT) can be related to each other.

**Proof of Corollary 1.** If the density is independent across goods then, using (30) in the Appendix and setting  $\gamma = 1$  yields

$$\frac{\partial^2 \pi}{\partial p_B \partial p_2} = 2f_1(p_B - p_2)(1 - F_2(p_2)) - f_1(p_B - p_2)f_2(p_2)p_2 + (p_B - p_2)f_1'(p_B - p_2)(1 - F_2(p_2)).$$

So implementing (X-sell) of Proposition 2 yields that dynamic pricing increases profits from the best mixed bundling prices  $(p_1, p_2, p_B)$  if:

$$-2f_1(p_B - p_2) + f_1(p_B - p_2) \frac{f_2(p_2)}{1 - F_2(p_2)} p_2 - (p_B - p_2) f_1'(p_B - p_2) > 0. \quad (22)$$

Now note that at the best mixed bundling prices we have  $\partial\pi_2/\partial p_2 + \partial\pi_B/\partial p_2 = 0$ . The derivative  $\partial\pi_B/\partial p_2$  is given in (29). The derivative  $\partial\pi_2/\partial p_2$  follows from (16). Combining, the best mixed bundling prices satisfy

$$p_2 \frac{f_2(p_2)}{1 - F_2(p_2)} - (p_B - p_2) \frac{f_1(p_B - p_2)}{F_1(p_B - p_2)} = 1. \quad (23)$$

Using (23) the condition for profitable dynamic pricing on the cross-sell (22) becomes

$$1 + (p_B - p_2) \frac{f_1(p_B - p_2)}{F_1(p_B - p_2)} - (p_B - p_2) \frac{f_1'(p_B - p_2)}{f_1(p_B - p_2)} > 2. \quad (24)$$

Now note that if  $x_1 f_1(x_1)/F_1(x_1)$  is declining in  $x_1$  then

$$x_1 \frac{f_1'(x_1)}{f_1(x_1)} - x_1 \frac{f_1(x_1)}{F_1(x_1)} < -1.$$

Setting  $x_1 = p_B - p_2$  yields the result. ■

We now apply Corollary 1 to some prominently studied settings.

### 6.2.1 The tilted uniform

Aguilera and Morin (2008) show numerically, and Manelli and Vincent (2006) theoretically, that the seller optimally uses only mixed bundling, and so no delay or randomness in delivery, when consumers' valuations are uniform on the square  $[0, 1]^2$ . We demonstrate that, maintaining the support on  $[0, 1]^2$ , the uniform distribution on a square is a knife-edge case. Even small perturbations in the density function yield that profits can be improved by dynamic pricing on the cross-sell.

Consider a distribution we refer to as the *tilted uniform* which is formed by tilting the graph of the uniform distribution as shown in Figure 4. More formally consider consumers with valuations on support  $[0, 1]^2$  with  $f_2 \equiv 1$  (uniform) and  $f_1(x_1) = 1 + \beta - 2\beta x_1$  for  $\beta \in [0, 1]$  implying  $F_1(x_1) = x_1(1 + \beta - \beta x_1)$ . If  $\beta = 0$  then we have the standard uniform case. If  $\beta > 0$  then we have a tilted uniform in which there is a downwards sloping density for good 1.

**Proposition 3** *Suppose that consumers are distributed according to the tilted uniform with parameter  $\beta > 0$ , then dynamic pricing on the cross-sell increases profits.*

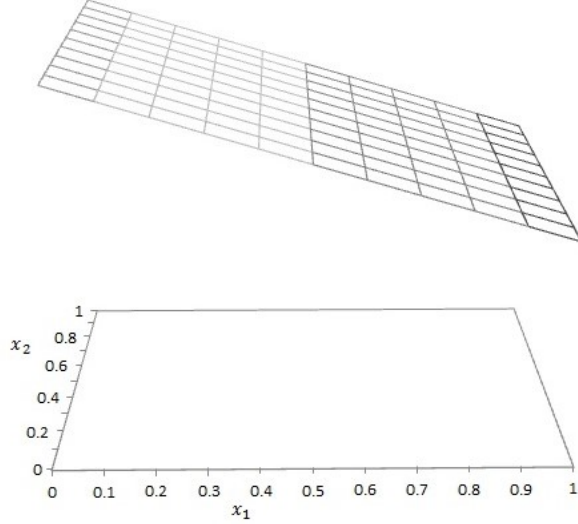


Figure 4: The tilted uniform

Notes: The surface plot is of the density function of consumers' valuations when tastes are distributed according to the tilted uniform supported on  $[0, 1]^2$ .

**Proof.** Note that

$$\frac{x_1 f_1(x_1)}{F_1(x_1)} = \frac{1 + \beta - 2\beta x_1}{1 + \beta - \beta x_1} = 2 - \frac{1 + \beta}{1 + \beta - \beta x_1}. \quad (25)$$

Since  $\beta > 0$  this function is decreasing in  $x_1$  and so the result follows from Corollary 1 if ABS is satisfied. It is known that the best bundle prices for the uniform distribution are uniquely defined and satisfy ABS.<sup>17</sup> ABS is therefore satisfied for small  $\beta$  by continuity of the extremal points inherited from the continuity of the profit function in its arguments. Numerical calculations confirm that ABS is satisfied for  $\beta \in [0, 1]$ . ■

Proposition 3 could not have been deduced by using Salant (1989) to convert the problem to one of optimising over lotteries and applying the extant literature. As we noted above, Manelli and Vincent (2006) establish sufficient conditions for mixed bundling prices to be fully optimal; the conditions are not proved necessary. Manelli and Vincent (2007) determine that there exist distributions for which lotteries (and so dynamic pricing) is optimal, but it does not confirm whether this is the case for the tilted uniform. In principle the solution method of Daskalakis et al. (2017) could be applied to the mixed bundling prices to establish that the prices cannot be optimal – however how the prices can be improved upon would require solving a transportation problem derived from the dual formulation. Our Proposition 2 is direct and constructive in demonstrating that dynamic pricing on the cross-sell is a more profitable strategy than the mixed bundling prices.

<sup>17</sup>Evaluating the first order conditions for the uniform distribution on  $[0, 1]^2$  yields  $p_1 = p_2 = 2/3$  and  $p_B = 2(2 - \sqrt{1/2})/3$ .

### 6.2.2 The shifted uniform

Inspection of Figure 1 demonstrates that dynamic pricing on the cross-sell expands the population of buyers slightly. If the consumer density function is supported only at very high valuations then this is likely to be important:

**Proposition 4** *Suppose that consumers are distributed according to the uniform distribution supported on  $[\underline{x}, \underline{x} + 1]^2$ . If ABS is satisfied at optimal bundling prices, then dynamic pricing increases profits for any strictly shifted uniform distribution of consumer valuations ( $\underline{x} > 0$ ).*

A related result to this in the setting of optimal pricing over lotteries has been confirmed via direct analytical calculation by Pavlov (2011a). Our method is more direct. Once again note that ABS is satisfied for small  $\underline{x}$  by continuity of the profit function:

**Proof.** Let  $F_1(x_1) = x_1 - \underline{x}$  on  $[\underline{x}, \underline{x} + 1]$ , then  $\frac{x_1 f_1(x_1)}{F_1(x_1)} = \frac{x_1}{x_1 - \underline{x}} = 1 + \frac{\underline{x}}{x_1 - \underline{x}}$ , which is decreasing in  $x_1$  for  $\underline{x} > 0$ . The result follows by Corollary 1. ■

### 6.2.3 Sufficient conditions for and against delayed reductions with independent valuations

Conditions (RT), derived here, and (MV), proved in Manelli and Vincent (2006), offer a partial categorisation of the space of distributions according to whether delayed price reductions are more profitable than standard mixed bundling. Instead of testing an inevitably ad hoc list of distributions against both (RT) and (MV), in this section we determine a distribution function which exactly separates the two conditions.

A distribution is on the cusp of satisfying the sufficiency condition (RT) if it satisfies the saturated differential equation. This is the O.D.E. from (RT) derived by setting the inequality condition to equality. We search for a distribution function which satisfies the system of saturated differential equations generated by both (MV) and (RT):

$$\left( \frac{x f_i(x)}{F_i(x)} \right)' = 0 = \left( \frac{x f_i'(x)}{f_i(x)} \right). \quad (26)$$

This pair of O.D.E.s have the same solution: the Power function distribution  $F_i(x) = x^c$  for  $x \in (0, 1]$  with Power function density  $f_i(x) = c x^{c-1}$ , and any constant  $c > 1$ .<sup>18</sup> Hence the Power function distribution is on the cusp of satisfying the sufficient condition for delay to be optimal, and the sufficient condition for no delay to be optimal. An appropriate small deviation in the distribution would allow one or other condition to be triggered.

---

<sup>18</sup>The density function formed by taking the product of the Power function density for each component ( $f(x) = \prod_i c_i x_i^{c_i-1}$ ) satisfies all three of the (MV) maintained assumptions. See footnote 16.

We first demonstrate that for (RT) to hold, the cumulative distribution function  $F_i(x)$  for some  $i$  must be more log-concave than that of the Power function distribution. If  $F_i(x)$  is log-concave then by definition  $f_i(x)/F_i(x)$  is decreasing in  $x$ . However for (RT) to hold we require this effect to dominate the increasing function  $x$  and so  $F_i(x)$  must be sufficiently log-concave. The Power function distribution is the class of probability distributions which is just on the boundary of satisfying (RT), and so any probability distribution which is more log-concave at all values of  $x$  will satisfy (RT).

By contrast, for (MV) to hold, the density function must be less log-concave than the density of the Power function distribution. If  $f_i(x)$  is log-concave then by definition  $f'_i(x)/f_i(x)$  is decreasing in  $x$ . For the (MV) condition to be satisfied we therefore require the log-concavity of  $f_i(x)$  to be sufficiently mild that it is overwhelmed by the increasing function  $x$ . The Power function distribution is the class of probability distributions which is just on the boundary of satisfying (MV), and so any distribution whose density function is less log-concave at all values of  $x$  and for each  $i$  will satisfy (MV).

## 7 Discussion

### 7.1 Real-world examples of price reductions over time

Our work discusses the profitability of targeted price reductions based on purchase history. The marketing industry has noted that the potential to implement such an approach is facilitated by the advent of ‘big data’. McKinsey<sup>19</sup> reports that using purchase history to identify cross-sell opportunities and then running microcampaigns around those opportunities increased revenues by a factor of five in a logistic company they work with. This potential has created an industry of data-analytics companies who market the ability to personalise their clients’ price promotion activities.<sup>20</sup> And as such sales strategies become increasingly established, consumers may be expected to develop increasingly accurate rational expectations of the future price drops.

The promise of using customer history to drive sales has a long pedigree in marketing. Deighton and Blattberg (1991) report an early example from AT&T who sought to increase take up of the bundle of telephony and credit-cards by offering consumers who had bought telephony and not credit-cards targeted discounts on credit.

A specific example of intertemporal price discrimination on purchase history can be given by one of the authors who is a customer of the BT Group for his internet service. BT seeks to encourage this author to upgrade to the bundle and buy mobile telephony data

---

<sup>19</sup>“Unlocking the power of data in sales”, McKinsey, December 2016. Available at: <https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/unlocking-the-power-of-data-in-sales> Accessed 28th August 2018.

<sup>20</sup>See for example the analytics marketing of Dunnhumby available at <https://www.dunnhumby.com/4-ways-improve-retail-promotions-planning-predictive-analytics-0> Accessed 28th August 2018.

in addition to the internet. Two months after signing for the single component (internet), the cross-sell to 20GB of mobile data was marketed to this author through personal email communication at £20 a month. After a further month's delay the price for the upgrade had dropped to £15 a month.<sup>21</sup> To emphasize the targeted nature of this offer the BT Group wrote: *"Selected for you, you won't see us advertise these deals anywhere else ..."* These examples demonstrate that some sellers are using purchase history to price discriminate over time.

## 7.2 Welfare Implications

The introduction of a dynamic price on the cross-sell results in a change of consumption: some consumers who would buy the bundle in the absence of dynamic pricing on the cross-sell swap to buying one component good first and the second only after delay, lowering welfare. Others who would only have bought one component good in the absence of the dynamically priced cross-sell decide to purchase the second component, albeit after delay, increasing welfare.

If the seller holds mixed bundling prices constant and introduces a dynamically priced cross-sell offer then consumer surplus must rise by revealed preference; only those consumers who prefer the cross-sell will buy it. If the introduction of the cross-sell is profitable then welfare has risen as welfare is the sum of consumer and producer surplus.

If the firm re-optimises its whole menu of prices including the cross-sell, the welfare effects are more difficult to determine as solving for the fully optimal selling strategy in the continuous case is not generally tractable. We can make progress by considering the two-type case of Section 3.

Return therefore to the two-type case in which consumers  $a$  and  $b$  are not strongly ordered.<sup>22</sup> If the seller is restricted to not use intertemporal pricing then we noted there are only two possible optimal cases. The first is to deliver the bundle to both types of consumer; hence  $p_B = a_{12}$  and profits are  $\pi^{\text{serve all}} = a_{12}$ . Alternatively the seller can deliver the full bundle to type  $b$  consumers and sell type  $a$  consumers just good 2. This requires prices  $p_2 = a_2$ ,  $p_B = b_{12}$ ; and yields profit  $\pi^{\text{discriminate}} = \lambda^a a_2 + \lambda^b b_{12}$ . Comparing the two cases it is immediate that

$$\pi^{\text{serve all}} > \pi^{\text{discriminate}} \iff \lambda^b < \frac{a_{12} - a_2}{b_{12} - a_2}. \quad (27)$$

Now we can study how the optimal selling strategy including dynamic pricing (using Appendix A.1 to strengthen Proposition 1) compares to the best prices without intertemporal price discrimination; and this is done in Figure 5. We see that  $\lambda^b$ , the proportion of

<sup>21</sup>The lower price has remained available via targeted email communications through to the time of writing.

<sup>22</sup>So that (1), (2) and (4) hold.



type  $b$  consumers, is separated into three regions:  $\{W_0, W_-, W^+\}$ . In region  $W_0$  dynamic pricing is suboptimal, and so welfare is unaffected by the possibility of its use. It is perhaps expected that if the proportion of type  $b$  consumers is large enough then without delay type  $a$  would not be served the bundle, but if delay is used then  $a$  will receive the bundle eventually and so welfare rises. This is region  $W^+$ . However, since  $a_2 > b_2$ , we see that there always exists a second region:  $W_-$  in Figure 5. Here welfare declines from the introduction of intertemporal price discrimination. The type  $a$  consumers in region  $W_-$  are numerous enough to receive the bundle and keep prices down if delay is not possible; but if delay is possible then the seller would lower the welfare type  $a$  receive so as to increase the prices to type  $b$ .

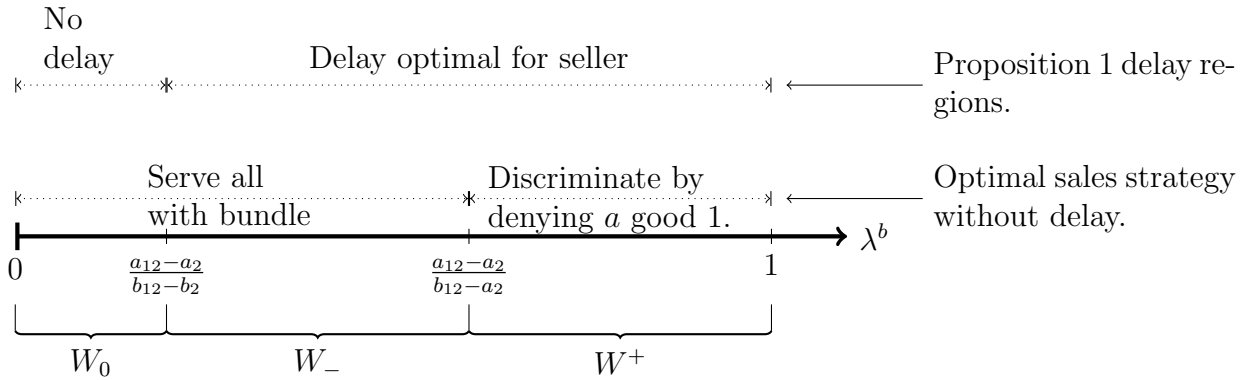


Figure 5: Welfare effects of dynamic pricing on the cross-sell

Notes: Recall  $\lambda^b$  is the proportion of  $b$  consumers in the two-type case of Section 3. In region  $W_-$  welfare is reduced by dynamic pricing. In region  $W^+$  welfare is increased by dynamic pricing. In region  $W_0$  welfare is unaffected by the potential for dynamic pricing as delay is sub-optimal.

### 7.3 The ABS assumption: All Bundles Sold

The ABS assumption holds that at the posted bundling prices, non-zero volumes of all product combinations are sold. In this section we explore under what circumstances we can guarantee that ABS holds.

Consider the case of  $\gamma = 1$  in which consumers see the goods as neither complements nor substitutes. In this case setting constant marginal costs equal to zero is without loss of generality: any arbitrary set of valuations can be translated by  $(-c_1, -c_2)$ . The valuations  $x$  of the model would then denote valuations above cost, and the prices set by the seller would be measuring the margins required. If consumers have free disposal then allowing for positive marginal costs of production would see consumers having valuations supported on  $x \in [-c_1, \infty) \times [-c_2, \infty)$ . We can in fact be still more general:

**Lemma 1** *Suppose that  $f$  is supported on  $[\underline{x}_1, \bar{x}_1] \times [\underline{x}_2, \bar{x}_2]$  with  $\underline{x}_i < 0 < \bar{x}_i$ , and  $\bar{x}_i$  may be infinite ( $i \in \{1, 2\}$ ). Then if the goods are neither complements nor substitutes ( $\gamma = 1$ ) ABS must be satisfied at optimal mixed bundling prices.*

Thus suppose the density  $f$  satisfies the support property of Lemma 1 then, by continuity of the profit function, if ABS is satisfied at  $\gamma = 1$  then it is satisfied for  $\gamma$  close to 1. And so Lemma 1 informs us that in general at *optimal* bundling prices, Proposition 2 applies. **Proof.** We first rule out pure bundling as optimal. Suppose otherwise so that the seller maximises profits by offering only the bundle at price  $p_B$  such that  $p_B < \bar{x}_2$  then only consumers with valuations such that  $x_1 + x_2 \geq p_B$  will purchase. If the seller were to introduce good 2 at a price of  $p_2 = p_B$  then she will increase her profits. This follows as good 2 sales generate the same margin over cost as the bundle (equal to  $p_B$ ), and some new consumers are served with good 2; in particular consumers in the triangle  $x_1 \in [\underline{x}_1, 0)$  and  $x_2 \in [p_B, p_B - x_1]$ . This contradicts the assumption of optimality.

Suppose instead the seller maximises her profits by selling just the bundle at price  $p_B$  with  $p_B \geq \bar{x}_2$ . Consider introducing a good 2 offer at a price of  $p_2 = \bar{x}_2 - \varepsilon$ . The bundle consumers who swap to the component lie in a triangle in the support with area order  $O(\varepsilon^2)$ .<sup>23</sup> While the set of consumers  $\{x : \underline{x}_1 < x_1 < p_B - \bar{x}_2 \text{ and } \bar{x}_2 - \varepsilon < x_2 < \bar{x}_2\}$  begin to buy and this is of order  $O(\varepsilon)$ . Hence for small  $\varepsilon$  this is profitable and so we have a contradiction.

Note that the same arguments ensure a contradiction if the bundle is sold with one but not the other component good. Next observe that if both components are being sold then the shape of the support guarantees that some consumers buy both products.

The final case to rule out is that only one component good, say good 2, is being sold at a price of  $p_2$ . But this cannot be optimal as profits could be increased by also selling good 1 at price  $p_1 = \bar{x}_1 - \varepsilon$  and setting  $p_B = p_1 + p_2$ . This increases consumers without losing any purchasers of good 2 (some upgrade to the bundle) and so raises profits – a contradiction. ■

Lemma 1 demonstrates that if there are some consumers who value one of the components below cost, then ABS is typically satisfied.

## 8 Conclusion

This paper has studied the two-product monopolist's pricing problem. The objective of the seller is to identify how to profitably alter prices over time as a function of the basket of products bought. Bundling and time-varying prices, including conditional on the purchase history, are common in many markets: cable companies bundling different TV channels and offering reductions to upgrade after some delay; internet retailers offering bundles at consumer specific prices which vary over time; supermarkets targeting coupons at shoppers with particular purchase histories.

We have shown that if the seller is serving sub-groups which are not strongly ordered,

---

<sup>23</sup>The consumers who swap from the bundle to good 2 lie in the triangle with coordinates  $(p_B - \bar{x}_2, \bar{x}_2)(p_B - \bar{x}_2 + \varepsilon, \bar{x}_2)(p_B - \bar{x}_2 + \varepsilon, \bar{x}_2 - \varepsilon)$ .

so that they differ in the products they value most highly, then dynamic pricing on the cross-sell is optimal if the consumer group which values the bundle most is sufficiently numerous. The profit maximising seller will use dynamic prices to delay the supply of a component of the bundle to a later date for one sub-group. This practice has ambiguous outcomes for welfare which depend upon the proportion of each consumer group in the population.

In the continuous setting we establish a sufficient condition for dynamic pricing on the cross-sell to be more profitable than mixed bundling prices: this condition is given by the sign of the cross-partial of the profit function with respect to the bundle and a component good price, and the condition holds true even if consumers have complementarities or substitutabilities in demand.

Sellers rarely sell only one good, and so extending the literature on dynamic pricing to multiple goods is important. Our work has studied the classical problem of a seller without binding capacity constraints. This setting fits well, in our view, to sectors where restocking is rapid, such as many types of retail (online, fast-moving-consumer-goods), or to sectors which do not involve the sale of a physical good (e.g. TV channels). How these sectors can maximise their profits across multiple time periods remains an open question.

## A Technical Proofs

### A.1 Extension of Proposition 1

In this section we extend Proposition 1 to demonstrate that the optimal sales strategy involves price discriminating over time if and only if (10) holds; and in this case is captured by (11). We first establish the general optimisation problem (7) and then proceed to solve this subject to  $(IR_a)$  and  $(IC_b)$  only, before checking the other constraints.

Re-working (3) for the case of  $t_1^a \leq t_2^a$  we establish that for general delivery times we have  $U(q_1^a, q_2^a; a) = q_1^a a_1 + q_2^a a_2 + \min(q_1^a, q_2^a)(a_{12} - a_1 - a_2) - p^a$ . Proceed as above to write down the individual rationality constraint of  $a$  and the incentive compatibility constraint of  $b$ , with general  $q_i^a, q_i^b$  for  $i \in \{1, 2\}$ . Then raising prices so that both of these constraints are satisfied with equality yields:

$$\begin{aligned} p^a &= q_1^a a_1 + q_2^a a_2 + \min(q_1^a, q_2^a)(a_{12} - a_1 - a_2) \\ p^b &= q_1^b b_1 + q_2^b b_2 + \min(q_1^b, q_2^b)(b_{12} - b_1 - b_2) \\ &\quad - q_1^a b_1 - q_2^a b_2 - \min(q_1^a, q_2^a)(b_{12} - b_1 - b_2) \\ &\quad + q_1^a a_1 + q_2^a a_2 + \min(q_1^a, q_2^a)(a_{12} - a_1 - a_2). \end{aligned}$$

As before  $\Pi = \lambda^a p^a + \lambda^b p^b$ . Now observe that if  $q_2^b \geq q_1^b$  then the profit function is increasing in  $q_2^b$ , and also in  $q_1^b$ ; the same follows if  $q_2^b \leq q_1^b$ . So optimality requires  $q_2^b = q_1^b = 1$ . Now

we show that optimality requires  $q_1^a \leq q_2^a$  so  $a$  receives good 2 first. Pick any  $s < t$  in  $[0, 1]$  and set  $q_1^a = s, q_2^a = t$ . The profit in this case, given  $q_2^b = q_1^b = 1$ , can be written as  $\Pi^{st} = s(a_{12} - a_2) + ta_2 + \lambda^b [b_{12} - s(b_{12} - b_2) - tb_2]$ . Reversing the assignment of  $s$  and  $t$  we can establish  $\Pi^{ts}$ . We then have that  $\Pi^{st} - \Pi^{ts} = (t - s) [(a_2 - a_1) - \lambda^b(b_2 - b_1)]$ . But  $(2) + (4) \Rightarrow a_2 - a_1 > b_2 - b_1$ . Hence  $\Pi^{st} > \Pi^{ts}$  as claimed. Therefore given  $q_1^a \leq q_2^a$  the coefficient of  $q_2^a$  in the profit function is  $a_2 - \lambda^b b_2 > 0$  by (4), and so setting  $q_2^a = 1$  would be optimal. Thus profit maximisation under (2) and (4) requires (6) to hold. The rest of the proof then proceeds as described above.

## A.2 Proof of Proposition 2

**Lemma 2** *Using  $\Delta\Pi$  given in (17),  $\lim_{\delta \searrow 0} \frac{d^2}{d\delta^2} \Delta\Pi > 0$  if and only if (X-sell) holds.*

**Proof.** Taking the first derivative of  $\Delta\Pi$  with respect to  $\delta$  yields

$$\begin{aligned} & \frac{d}{d\delta} [\Delta\Pi] \\ = & -q_1 \left\{ \int_{x_2=p_2-\delta\frac{q_1}{1-q_1}}^{p_2} \int_{x_1=-x_2\left(\frac{1-q_1+q_1\gamma}{q_1\gamma}\right)+\frac{p_B}{\gamma}+\left(\frac{1-q_1}{q_1\gamma}\right)p_2-\frac{\delta}{\gamma}}^{(\frac{1-\gamma}{\gamma})x_2+\frac{p_B-p_2}{\gamma}+\frac{\delta}{\gamma}\frac{q_1}{1-q_1}} dF \right. \\ & \quad \left. + \int_{x_2=p_2}^{\infty} \int_{x_1=(\frac{1-\gamma}{\gamma})x_2+\frac{p_B-p_2}{\gamma}-\frac{\delta}{\gamma}}^{(\frac{1-\gamma}{\gamma})x_2+\frac{p_B-p_2}{\gamma}+\frac{\delta}{\gamma}\frac{q_1}{1-q_1}} dF \right\} \\ & + [q_1 p_B + (1 - q_1) p_2 - q_1 \delta] \left\{ \begin{aligned} & \frac{1}{\gamma} \int_{x_2=p_2-\delta\frac{q_1}{1-q_1}}^{p_2} f \left( -x_2 \left( \frac{1-q_1+q_1\gamma}{q_1\gamma} \right) + \frac{p_B}{\gamma} + \left( \frac{1-q_1}{q_1\gamma} \right) p_2 - \frac{\delta}{\gamma}, x_2 \right) dx_2 \\ & + \frac{1}{\gamma} \frac{q_1}{1-q_1} \int_{x_2=p_2-\delta\frac{q_1}{1-q_1}}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B-p_2}{\gamma} + \frac{\delta}{\gamma} \frac{q_1}{1-q_1}, x_2 \right) dx_2 \\ & + \frac{1}{\gamma} \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B-p_2}{\gamma} - \frac{\delta}{\gamma}, x_2 \right) dx_2 \end{aligned} \right\} \\ & - p_2 \frac{1}{\gamma} \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B-p_2}{\gamma} - \frac{\delta}{\gamma}, x_2 \right) dx_2 \\ & - p_B \left\{ \frac{1}{\gamma} \frac{q_1}{1-q_1} \int_{x_2=p_2-\delta\frac{q_1}{1-q_1}}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B-p_2}{\gamma} + \frac{\delta}{\gamma} \frac{q_1}{1-q_1}, x_2 \right) dx_2 \right\} . \end{aligned}$$

This vanishes at  $\delta = 0$ . We determine  $\frac{d^2}{d\delta^2} [\Delta\Pi]_{\delta=0}$  and simplify to yield:

$$\begin{aligned}
\frac{d^2}{d\delta^2} [\Delta\Pi]_{\delta=0} &= -2\frac{1}{\gamma} \left( \frac{q_1}{1-q_1} \right) \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\
&\quad + [q_1 p_B + (1-q_1) p_2] \left\{ \begin{aligned} &\frac{1}{\gamma} \frac{q_1}{1-q_1} f \left( \frac{p_B}{\gamma} - p_2, p_2 \right) \\ &+ \frac{1}{\gamma} \left( \frac{q_1}{1-q_1} \right)^2 f \left( \frac{p_B}{\gamma} - p_2, p_2 \right) \\ &+ \left( \frac{1}{\gamma} \frac{q_1}{1-q_1} \right)^2 \int_{x_2=p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\ &- \left( \frac{1}{\gamma} \right)^2 \int_{x_2=p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \end{aligned} \right\} \\
&\quad + p_2 \left( \frac{1}{\gamma} \right)^2 \int_{x_2=p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\
&\quad - p_B \left\{ \left( \frac{1}{\gamma} \frac{q_1}{1-q_1} \right)^2 \int_{x_2=p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \right\} \\
&\quad - p_B \left\{ \frac{1}{\gamma} \left( \frac{q_1}{1-q_1} \right)^2 f \left( \frac{p_B}{\gamma} - p_2, p_2 \right) \right\} .
\end{aligned}$$

So collecting terms:

$$\begin{aligned}
\frac{d^2}{d\delta^2} [\Delta\Pi]_{\delta=0} &= -2\frac{1}{\gamma} \left( \frac{q_1}{1-q_1} \right) \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\
&\quad + f \left( \frac{p_B}{\gamma} - p_2, p_2 \right) \frac{1}{\gamma} \left( \begin{aligned} &[q_1 p_B + (1-q_1) p_2] \left[ \left( \frac{q_1}{1-q_1} \right)^2 + \frac{q_1}{1-q_1} \right] \\ &- p_B \left( \frac{q_1}{1-q_1} \right)^2 \end{aligned} \right) \\
&\quad - \frac{1}{\gamma^2} \int_{p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \left( \begin{aligned} &[q_1 p_B + (1-q_1) p_2] \left( 1 - \left( \frac{q_1}{1-q_1} \right)^2 \right) \\ &- p_2 + p_B \left( \frac{q_1}{1-q_1} \right)^2 \end{aligned} \right)
\end{aligned}$$

which simplifies again to give:

$$\frac{d^2}{d\delta^2} [\Delta\Pi]_{\delta=0} = \frac{q_1}{1-q_1} \frac{1}{\gamma} \left\{ \begin{aligned} &-2 \int_{p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\ &+ f \left( \frac{p_B}{\gamma} - p_2, p_2 \right) p_2 \\ &- (p_B - p_2) \frac{1}{\gamma} \int_{p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \end{aligned} \right\} . \quad (28)$$

We now turn to the profit function for the seller in the absence of dynamic pricing. Using (16):

$$\begin{aligned}
\frac{\partial \pi_2}{\partial p_B} &= \frac{1}{\gamma} p_2 \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\
\frac{\partial^2 \pi_2}{\partial p_B \partial p_2} &= \frac{1}{\gamma} \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 - \frac{1}{\gamma} p_2 f \left( \frac{p_B}{\gamma} - p_2, p_2 \right) \\
&\quad - \frac{1}{\gamma^2} p_2 \int_{x_2=p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 .
\end{aligned}$$

Similarly using (16)

$$\frac{\partial \pi_B}{\partial p_2} = \frac{1}{\gamma} p_B \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2. \quad (29)$$

Hence

$$\begin{aligned} \frac{\partial^2 \pi_B}{\partial p_B \partial p_2} &= \frac{1}{\gamma} \int_{x_2=p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\ &\quad + \frac{1}{\gamma^2} p_B \int_{x_2=p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2. \end{aligned}$$

Combining it follows that

$$\frac{\partial^2 \pi}{\partial p_B \partial p_2} = \frac{1}{\gamma} \left\{ \begin{aligned} &2 \int_{p_2}^{\infty} f \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \\ &- p_2 f \left( \frac{p_B}{\gamma} - p_2, p_2 \right) \\ &+ (p_B - p_2) \frac{1}{\gamma} \int_{p_2}^{\infty} \frac{\partial f}{\partial x_1} \left( \left( \frac{1-\gamma}{\gamma} \right) x_2 + \frac{p_B - p_2}{\gamma}, x_2 \right) dx_2 \end{aligned} \right\}, \quad (30)$$

and so we have

$$\frac{d^2}{d\delta^2} [\Delta \Pi]_{\delta=0} = -\frac{q_1}{1 - q_1} \frac{\partial^2 \pi}{\partial p_B \partial p_2}.$$

The result follows. ■

## References

- Adams, W. J. and J. L. Yellen (1976). Commodity bundling and the burden of monopoly. *The Quarterly Journal of Economics*, 475–498.
- Aguilera, N. E. and P. Morin (2008). Approximating optimization problems over convex functions. *Numerische Mathematik* 111(1), 1–34.
- Anderson, E. T. and J. D. Dana Jr (2009). When is price discrimination profitable? *Management Science* 55(6), 980–989.
- Armstrong, M. (2013). A more general theory of commodity bundling. *Journal of Economic Theory* 148(2), 448–472.
- Bakos, Y. and E. Brynjolfsson (1999). Bundling information goods: Pricing, profits, and efficiency. *Management Science* 45(12), 1613–1630.
- Besbes, O. and I. Lobel (2015). Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science* 61(1), 92–110.
- Blattberg, R. C., B.-D. Kim, and S. A. Neslin (2008). *Database Marketing Analyzing and Managing Customers*. Springer-Verlag New York.
- Board, S. and A. Skrzypacz (2016). Revenue management with forward-looking buyers. *Journal of Political Economy* 124(4), 1046–1087.

- Coase, R. H. (1972). Durability and monopoly. *The Journal of Law and Economics* 15(1), 143–149.
- Daskalakis, C., A. Deckelbaum, and C. Tzamos (2017). Strong duality for a multiple-good monopolist. *Econometrica* 85, 735–767.
- Deighton, J. A. and R. C. Blattberg (1991). Interactive marketing: exploiting the age of addressability. *Sloan management review* 33(1), 5–14.
- Deneckere, R. J. and P. R. McAfee (1996). Damaged goods. *Journal of Economics & Management Strategy* 5(2), 149–174.
- Elmaghraby, W., A. Gülcü, and P. Keskinocak (2008). Designing optimal preannounced markdowns in the presence of rational customers with multiunit demands. *Manufacturing & Service Operations Management* 10(1), 126–148.
- Ferrell, O. C. and M. Hartline (2012). *Marketing strategy, text and cases*. Nelson Education.
- Gul, F., H. Sonnenschein, and R. Wilson (1986). Foundations of dynamic monopoly and the coase conjecture. *Journal of Economic Theory* 39(1), 155–190.
- Hart, S. and N. Nisan (2017). The menu-size complexity of auctions. *arXiv preprint arXiv:1304.6116*.
- Hart, S. and P. J. Reny (2015). Maximal revenue with multiple goods: Nonmonotonicity and other observations. *Theoretical Economics* 10(3), 893–922.
- Johnson, J. P. and D. P. Myatt (2003). Multiproduct quality competition: Fighting brands and product line pruning. *American Economic Review* 93(3), 748–774.
- Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright (1998). Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on optimization* 9(1), 112–147.
- Manelli, A. M. and D. R. Vincent (2006). Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory* 127(1), 1–35.
- Manelli, A. M. and D. R. Vincent (2007). Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory* 137(1), 153–185.
- McAfee, R. P., J. McMillan, and M. D. Whinston (1989). Multiproduct monopoly, commodity bundling, and correlation of values. *The Quarterly Journal of Economics* 104(2), 371–383.
- Menicucci, D., S. Hurkens, and D.-S. Jeon (2015). On the optimality of pure bundling for a monopolist. *Journal of Mathematical Economics* 60, 33–42.
- Pavlov, G. (2011a). Optimal mechanism for selling two goods. *The BE Journal of Theoretical Economics* 11(1).
- Pavlov, G. (2011b). A property of solutions to linear monopoly problems. *The BE Journal of Theoretical Economics* 11(1).

- Prinzie, A. and D. Van den Poel (2005). Constrained optimization of data-mining problems to improve model performance: A direct-marketing application. *Expert Systems with Applications* 29(3), 630–640.
- Pycia, M. (2006). Stochastic vs deterministic mechanisms in multidimensional screening. Technical report, Technical report, MIT.[893, 908, 912].
- Reutterer, T., A. Mild, M. Natter, and A. Taudes (2006). A dynamic segmentation approach for targeting and customizing direct marketing campaigns. *Journal of interactive Marketing* 20(3-4), 43–57.
- Riley, J. and R. Zeckhauser (1983). Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics* 98(2), 267–289.
- Rochet, J.-C. and P. Choné (1998). Ironing, sweeping, and multidimensional screening. *Econometrica*, 783–826.
- Rossi, P. E., R. E. McCulloch, and G. M. Allenby (1996). The value of purchase history data in target marketing. *Marketing Science* 15(4), 321–340.
- Salant, S. W. (1989). When is inducing self-selection suboptimal for a monopolist? *The Quarterly Journal of Economics* 104(2), 391–397.
- Schmalensee, R. (1984). Gaussian demand and commodity bundling. *Journal of Business*, S211–S230.
- Skreta, V. (2006). Sequentially optimal mechanisms. *The Review of Economic Studies* 73(4), 1085–1111.
- Sobel, J. (1991). Durable goods monopoly with entry of new consumers. *Econometrica*, 1455–1485.
- Stokey, N. L. (1979). Intertemporal price discrimination. *The Quarterly Journal of Economics*, 355–371.
- Talluri, K. T. and G. J. Van Ryzin (2006). *The theory and practice of revenue management*, Volume 68. Springer Science & Business Media.
- Tang, P. and Z. Wang (2017). Optimal mechanisms with simple menus. *Journal of Mathematical Economics* 69, 54–70.
- Taylor, C. R. (2004). Consumer privacy and the market for customer information. *RAND Journal of Economics*, 631–650.
- Thanassoulis, J. (2004). Haggling over substitutes. *Journal of Economic Theory* 117(2), 217–245.
- Venkatesh, R. and W. A. Kamakura (2003). Optimal bundling and pricing under a monopoly: Contrasting complements and substitutes from independently valued products. *Journal of Business* 76(2), 211–231.
- Zhang, J. and M. Wedel (2009). The effectiveness of customized promotions in online and offline stores. *Journal of marketing research* 46(2), 190–206.